DOCUMENT RESUME

ED 116 604 IR 002 467

AUTHOR Bruell, Jan H.
TITLE FIT: A New Computer-Scorable Test.
INSTITUTION Texas Univ., Austin. Project C-BE.
SPONS AGENCY National Science Foundation, Washington, D.C.
REPORT NO EP-5-1-25-73
PUB DATE 25 Jan 73
NOTE 25p.; For related documents see IR 002 463 and 464

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Computer Assisted Instruction; *Computer Programs; Test Construction; *Tests
IDENTIFIERS Computer Generated Tests; Computer Scorable Tests; Project C BE

ABSTRACT

Fill in tests (FIT) are composed and printed by the computer; the tests are machine scorable as well. Two-digit code numbers are assigned to words and numbers in a vocabulary or answers. FIT also permits the use of questions which demand student-constructed responses. FIT is a recognition test, but given a sufficient vocabulary list, it can become a recall-type test. (CH)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED116604

FIT - A NEW COMPUTER-SCORABLE TEST

EP-5/1/25/73

Jan H. Bruell
Department of Psychology
The University of Texas at Austin
Austin, Texas 78712

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IR 002 467

*This study was supported in part by PROJECT C-BE under Grant GY-9340, "The Use of Computer-Based Teaching Techniques in Undergraduate Science and Engineering Education," from the National Science Foundation to the University of Texas at Austin, Drs. John J. Allan and J. J. Lagowski, Co-Directors.

TABLE OF CONTENTS

		PAGE
PART I	FIT - A NEW COMPUTER SCORABLE TEST	
I.	INTRODUCTION	1
	PRODUCTION OF TESTS	2
	SCORING OF TESTS	6
	ADDENDA	8
	CONCLUSION	11
PART II	FORMS AND READERS FOR FIT	17

FIT - A NEW COMPUTER-SCORABLE TEST*

Jan H. Bruell
The University of Texas at Austin

I. INTRODUCTION

FIT, for Fill-In Test, is the name for (1) a new machine scorable test and (2) the computer program that produces and scores FIT tests. In its most simple form, FIT calls for the completion of statements such as "The sun rises in the _____," or "The German word for dog is _____," or "8 + 8 = _____."

What makes the test machine scorable is the fact that the student fills code numbers rather than words into the blank spaces of the text. Machine readable forms, similar to those used for multiple choice tests, are employed. Customarily, the answer to a multiple choice question is an a, b, c, d, or e; the answer to a FIT question is a two digit code number. Multiple choice forms provide for every question a row or column containing the letters a to e; FIT forms reserve per question two adjacent columns, each imprinted with the numbers 0 to 9. Thus, by cancelling the appropriate digits, the student can indicate any code number between 00 and 99.

A hypothetical FIT test is shown in Figure 1. The student has to fill in the blanks indicated by brackets in the text. The first blank appears

in the statement "JANUARY HAS XX (01) DAYS." Obviously, the student must replace the XX with a 31, and he, therefore, cancels on his answer sheet the digits 3 and 1 in the double column provided for question 1. The second blank appears in "THE SHORTEST MONTH OF THE YEAR IS (02)." The student scans the vocabulary list at the left side of his test and finds the entry "16 FEBRUARY." The 16 preceding the name of the month is the code number the student enters on his answer sheet. For one more example turn to the last test question: "NAME A SUMMER MONTH WITH 30 DAYS (05)." Since June and September have 30 days, the student receives credit for responding with code number 57 or code number 31, numbers he finds in the vocabulary.

Though simple minded in conception, FIT has proven to be a surprisingly versatile and powerful test. We present it here as a possible alternative to the multiple choice test, the dominant test tool in American education.

II. PRODUCTION OF TESTS

FIT tests are composed and printed by computer. To see how this is done, let us examine the deck of cards that was used to print the test shown in Figure 1. The deck (Figure 2) consists of (1) data cards, i.e., cards providing the information needed for the generation of tests, and (2) instruction cards which control output operations.

A. Data

1. Text Cards

Text cards follow the READ TEXT instruction. The text, punched on 80-column tab cards exactly as shown in Figure 2, is divided into statements or paragraphs, each containing one or more blanks or questions. Paragraphs are identified by two-digit alphanumeric numbers punched in columns 1 and 2 of each text card. The text in Figure 2 consists of paragraphs AA, AB, AC, AD, and AE. Blanks within paragraphs are also marked by two digit numbers. By combining the identifiers of paragraphs and blanks, a unique number is provided for each blank in the text. For example, the unique number of the blank in paragraph AB is AB + 01 = AB01. Or, the unique numbers of the blanks in paragraph AA are AA01, AA02, and AA03.

2. Vocabulary Cards

The terms SEASON, FEBRUARY, LEAP, JUNE, and SEPTEMBER are needed to fill in blanks in the text. Each term is punched on a card and linked to the appropriate blank by punching the blank's unique number alongside. Thus, the term SEASON is linked to two blanks, AA02 and AA03, FEBRUARY is linked to AB01, and so on. The last two vocabulary cards in Figure 2 serve as distractors. Since neither JULY nor AUGUST complete statements of the test, these two cards have no unique numbers punched into them.

Observe that, on vocabulary cards, terms are not preceded by code numbers. The code numbers appearing on printed tests are generated by FIT which uses a special algorithm to translate the first ten characters of each term into a two digit code number.

3. Number Cards

The test shown in Figure 2 contains blanks calling for numerical answers. The numbers 12, 28, and 31 appear on the test, and a number card is punched for each one of them. These cards are linked to the appropriate blanks in the text the same way terms are linked to blanks. Thus, number card 12 is linked to AA 01, 28 to AC02, and 31 to AD01.

4. The Parameter Card

The parameter card tells the computer what tests to print. In our example, the computer is instructed to prepare for the printing of three versions (V = 06/08) of Test 12 (T = 12), and to number them 1206 to 1208. For each version, five questions (Q = 5) are to be drawn at random from the text with the proviso that paragraphs containing several questions are not to be broken up in the process. The computer is informed that 20 sets (S = 20), each set consisting of tests 1206 to 1208, are to be printed.

B. Output

The text, vocabulary, number, and parameter cards contain all the information needed for the production of many types of output. In this

brief description of FIT, however, we will limit discussion to three types: (1) sequence cards, (2) keys, and (3) printed tests.

1. Sequence Cards

The PUNCH SEQUENCE instruction (Figure 2) produces the sequence cards shown in Figure 3. Recall that, according to the parameter card, versions 1206, 1207, and 1208 of Test 12 were to be produced. Each version was to contain five questions drawn at random from the text. The questions selected for each version of the test are listed, by number, on the sequence cards reproduced in Figure 3. For example, the sequence card for version 1206 tells us that questions AD01, AE01, AC01, AC02, and AE01 will appear in this order on Test 1206. Test 1206 is shown in Figure 1, and it is easy to see that in fact it is composed of the questions just listed.

2. Keys

The set of key cards produced by the PUNCH KEYS command is shown in Figure 3. How were these cards generated and what information do they convey? Consider first cards 12 and 28. These are straight copies of number cards submitted by the user (see Figure 2). The cards establish links between numerical responses and the blanks or questions to which the responses pertain. Consider next key card 16. The number "16" is the computer-generated code number for FEBRUARY (Figure 1), and FEBRUARY is the correct response to question AB01 (Figure 2). These

two items of information are linked on key card 16: "16 AB01." Finally, consider key card 31. The number 31 is the numerical response to AD01. It is also the code for SEPTEMBER, and SEPTEMBER is the answer to AE01. Hence, both AD01 and AE01 appear on key 31.

3. Tests

Test 1206, already discussed in the introduction to this paper, was printed in response to the PRINT TESTS instruction. Two features of the printed test stand out: the terms in the vocabulary are alphabetized, although the vocabulary cards submitted by the user are not. And, the bracketed numbers appearing on the user's text cards are replaced by serial numbers. The numbers supplied by the user are unique, "absolute," question numbers; they do not depend on the position a question holds in a particular version of the test. By contrast, the numbers the student sees are "version bound;" they indicate the serial position of a question on a given test.

III. SCORING OF TESTS

We will explain how the computer goes about the scoring of FIT tests by showing how tests can be scored "by hand." For purposes of illustration, let us again refer to Test 1206, and proceed by preparing the following key:

Key for Test 1206

Serial number	(01)	(02)	(03)	(04)	(05)
Absolute number	AD01	AB01	AC01	AC02	AE01
Correct response	31	16	36	28	31,57

Line 2 of the key is a copy of sequence card 1206; the line contains the "absolute numbers" of the questions appearing in positions (1) through (5) of Test 1206. Line 3 contains the correct responses to these questions. They are read off the key cards of Figure 3. For example, AD01 appears on key card 31; thus code number 31 is the proper response to (01). Or, AB01 appears on key card 16; hence code number 16 is the correct answer to (02). Or, AE01 appears on key cards 31 and 57; therefore, either number is a satisfactory response to (05).

Using this key, we can score Test 1206. Suppose a student entered code numbers 31, 16, 36, 28, and 64 on his answer sheet. The first four responses match the key, the last one does not. Thus the student was 80 percent correct. Curiosity may prompt us to inquire what was on the student's mind when he gave his incorrect response to (05). The student's response was 64. Checking the vocabulary in Figure 1, we see that 64 is the code for JULY. The student was to name a summer month with 30 days, and he picked the wrong month.

An even faster procedure for the hand- scoring of tests makes direct use of the sequence and key cards of Figure 3 and obviates the preparation

of a special key. Consider again Test 1206 and the student whose response to (01) was 31. We check sequence card 1206 and find that AD01 appears in position (01) on the test. We then look at key card 31. If AD01 is on it, as it is, the student's response is correct. Or, the student responded to (05) with 64. We again look first at sequence card 1206, but, in this special case, unnecessarily. Turning to our key cards, we do not find a key 64; hence, the student's response is incorrect. FIT uses this simplified procedure for scoring. Sequence and key cards are stored in the memory of the computer. When a card with a student's responses to a given test is read in, the computer checks every response in the just described manner.

IV. ADDENDA

A. Code Numbers

Wishing to concentrate on essentials first, we have purposely postponed until now a more detailed discussion of code numbers. We have shown how code numbers are assigned to words, but have not shown how they are assigned to numbers. Recall our first example "JANUARY HAS XX (01) DAYS," in which the student replaced the XX with a 31. This is a simple case because a two digit integer is replaced by a two digit code number. But how do we deal with larger integers, or decimal numbers, or fractions, or exponents?

We ask the student to respond with those two digits of a number which on the test are indicated by an X. Consider, for example, possible results of computations on a test in statistics:

Result of Computation	Student is asked to report	Correct Answer
4328	XX	28
4328	XX--	43
23.56	XX.	23
23.56	X.X	35
23.56	.XX	56
3/4	X/X	34
5^{12}	XX	12

Using these conventions, we can ask questions about the results of any kind of computation. While only two digits of results can be reported, in most cases this is all that is needed to check the correctness of an answer.

Code numbers contribute in other ways to flexibility of test construction. For example, we have used them to point to items in tables and figures. In a test of neuroanatomy we show a picture of the brain. Various areas of the brain are numbered, and the student uses the numbers to respond to a series of questions such as:

IN FIGURE 8, POINT TO THE

RETICULAR FORMATION	XX (01)
CORPUS STRIATUM	XX (02)
THALAMUS	XX (03)

It goes without saying that the same figure can be used to ask such questions as "IN FIGURE 8, NUMBER 15 POINTS TO (04)." Here, the answer, e.g., "83 THALAMUS," is found in the vocabulary, and the student responds with code number 83.

One more important use of code numbers shall be mentioned: they permit one to ask multiple choice questions on FIT tests. Consider the following example:

(16) SELECT AND INDICATE BY NUMBER THE ODD ITEM:

01. THERMOMETER
02. BAROMETER
03. KILOMETER
04. SPEEDOMETER

Obviously, the correct response to question (16) is 03. This use of code numbers permits free intermixing of FIT and multiple choice questions.

B. Vocabulary Lists

The construction of vocabulary lists for FIT tests requires considerable care. The vocabulary shown in Figure 1 is too short and would hardly be a challenge to anyone. Two major criticisms have

been leveled against multiple choice tests: they are recognition rather than recall tests; among the few choices they offer, even a poorly prepared student can recognize the correct choice. They encourage guessing; by chance alone, a student will be right one out of four times. An overly short vocabulary list on a FIT test invites the same criticisms. In our experience, lists comprised of 30 to 40 terms appear to be just right. Guessing is all but eliminated, and recognition is made more difficult. In order to find it, the student must know what term he is looking for.

The grammatical form of a term, must not give it away. We use only the infinitive of verbs and the singular form of nouns, and where nouns, verbs, adverbs, and adjectives have the same root, we display them the way they might appear in a dictionary, e.g., CAUS-E, -AL, -ALITY.

Inclusion of distractors is often necessary. Suppose, for instance, a test asked for the name of the "father of psychoanalysis." If only Freud's name appeared in the vocabulary list, the name would stand out. Without distractors, the test would become even less valid than multiple choice tests sometimes are.

V. CONCLUSIONS

In conclusion, it may be well to compare the relative merits of multiple choice tests (MCT) and FIT, the fill-in tests described here.

(1) MCT and FIT can be composed and printed by computer. This is particularly important in self-paced classes where the ability to generate individualized tests is of the essence.

(2) MCT and FIT are machine scorable and thus well suited for self-paced classes and classes with large enrollments.

(3) MCT is a recognition test: the author provides the correct answer to a question, the student must recognize it. To the extent that the correct response to a FIT question is found in a vocabulary list, FIT is also a recognition test. But, with sufficiently long vocabulary lists, chance recognition of responses can be made difficult; under ideal conditions, FIT turns into a recall-type test.

(4) Also, FIT permits the use of questions which demand "constructed" responses: the student must generate the response rather than find it in a list. This is true of all questions involving computations. On this score, FIT surpasses the capabilities of MCT.

(5) For some types of problems (see second example on page 8), the format of multiple choice questions is better suited than the format of fill-in questions. This, however, does not prevent the use of FIT, since FIT permits the free intermixing of multiple choice and fill-in questions.

We conclude that FIT is similar to a multiple choice test in some respects, but that it surpasses it in others. We have found it to be a valuable tool for testing.

TEST 1206

05 AUGUST
16 FEBRUARY
64 JULY
57 JUNE
36 LEAP
83 SEASON
31 SEPTEMBER

JANUARY HAS XX [01] DAYS.

THE SHORTEST MONTH OF THE YEAR IS [02].

A YEAR IN WHICH THE SHORTEST MONTH HAS 29 DAYS IS CALLED [03] YEAR. IN OTHER YEARS THAT MONTH HAS XX [04] DAYS.

NAME A SUMMER MONTH WITH 30 DAYS [05].

Figure 1. Fill-in test printed by computer. Blanks to be filled are marked by brackets. Machine readable forms are used. Students fill in numbers, e.g., 31 for [01], 16 for [02], 36 for [03]. Numbers 16 and 36 are code numbers taken from vocabulary printed at left side of test.

```
READ TEXT
AA    THE YEAR HAS XX [01] MONTHS AND FOUR [02].
AA SPRING IS THE FIRST [03] OF THE YEAR.
AB    THE SHORTEST MONTH OF THE YEAR IS [01].
AC    A YEAR IN WHICH THE SHORTEST MONTH HAS 29
AC DAYS IS CALLED [01] YEAR.  IN OTHER YEARS
AC THAT MONTH HAS XX [02] DAYS.
AD    JANUARY HAS XX [01] DAYS.
AE    NAME A SUMMER MONTH WITH 30 DAYS [01].
End-of-Record card


READ VOCABULARY
  SEASON                AA02 AA03
  FEBRUARY              AB01
  LEAP                  AC01
  JUNE                  AE01
  SEPTEMBER             AE01
  JULY
  AUGUST
End-of-Record card


READ NUMBERS
12  AA01
28  AC02
31  AD01
End-of-Record card


PARAMETERS:  T=12 V=06/08 Q=5 S=20


PUNCH SEQUENCE
PUNCH KEYS
PRINT TESTS
STOP
```

Figure 2. Deck of cards that produced FIT test shown in Figure 1. See text for explanations.

Sequence cards:

```
1206  05  AD01  AB01  AC01  AC02  AE01
1207  05  AZ01  AA01  AA02  AA03  AB01
1208  05  AB01  AD01  AA01  AA02  AA03
..
```

Key cards:

```
12  AA01
16  AB01
28  AC02
31  AD01 AE01
36  AC01
57  AE01
83  AA02 AA03
..
```

Figure 3. Sequence and key cards produced by PUNCH SEQUENCE and PUNCH KEYS commands shown in Figure 2.

FORMS AND READERS FOR FIT

Efficient use of FIT requires special forms and special scanners for automatic recording of student responses. Several forms and associated scanning equipment are described below.

Forms for Optical Page Readers

Figure 1 shows a form produced by overprinting the general coding form DS2-970A of the Optical Scanning Corporation. Of course, a similar form could be custom printed. Unless large quantities are ordered, however, custom printed forms are quite expensive. The form shown in Figure 1 can be read by several available optical page readers. We had experience with three readers.

Optical Scanning Corporation, Model 100

This is a high speed optical scanner used primarily for large batch jobs. At the University of Texas the Measurement and Evaluation Center of the University uses it for the reading of multiple choice exams given on standard test forms in large classes. The unit reads the forms and either transcribes them to punched cards or scores them and produces punched score cards. Because of the special scoring required by FIT, we used Model 100 only for transcription. The unit is reliable but quite expensive. Few departments could afford placing it in a test room. Thus several steps are interposed between test taking and display of test scores to students: (1) test forms are mailed to a central transcription facility, (2) forms are transcribed to cards, (3) cards are read by computer in batch mode and scored, (4) lists with scores are printed by line printer and posted.

In passing, it should be mentioned that scoring of multiple choice tests given in selfpaced classes is not less time consuming. Model 100 is designed for batch scoring of exams. To score a test a key for the particular test must be read in first. The key readies the machine for scoring. If, as is the case in selfpaced classes, many kinds of test are given during the same testing session, the forms must be manually separated by test number, a key must be placed in front of each set of tests having the same number, and the scanner must be set up separately for each of these sub-batches. This becomes a time consuming and inefficient operation. To get around it, answer sheets for FIT tests entered in random order are transcribed (not scored!) to cards. These cards are then read by the computer and scored using a set of keys stored for this purpose in the memory of the computer. Thus the same route has to be taken for FIT and multiple data tests.

Optical Scanning Corporation, Model 12/17

This is a slower desk top edition of the scanner described above. The advantage of it is that it can be stationed in the classroom. At

the beginning of a testing session, the unit is programmed to read a particular test form. Then students using that form can read in their tests. A page image is transmitted to a central computer via telephone lines, the computer scores the test and returns the test score. The score is shown on a CRT or teletype. The unit is close to ideal with these restrictions: different forms cannot be intermixed; the unit must be preprogrammed for each form by advance scanning of appropriate template forms. The unit and maintenance contracts are quite expensive. In our experience, forms were often mangled by the reader and many repairs by a company repairman were necessary.

Decision OMR 650

This desk top optical reader reads Optical Scanning Corporation forms and can be used in precisely the same way the Opscan Model 12/17 is used. It can be set up in the classroom where it reads forms and transmits information by telephone to a computer. The unit is less expensive than Model 12/17 and, since it does not have an automatic feeding mechanism, it does not mutilate forms. Forms are placed singly on a glass screen, covered by a weight, and read. Thus feeding of forms is slow but sufficiently fast for the intended purpose. Unfortunately, we used one of the first units produced. The unit had a quite complex optical scanning system using oscillating mirrors that got out of adjustment easily. The scanner often read material it was not supposed to read (e.g., social security numbers entered by students above the field to be scanned) and we never could get it to read without error. Support from the company producing the unit was hard to obtain. Local service was unavailable and after much frustration, we gave up. But, if newer models of the machine are better, and the company builds up a service organization, this unit is worth considering for classroom use.

Forms for IBM Mark Sense Reader

The first widely used system for entering hand-written information into a computer was developed by IBM. It consisted of tabulating cards on which information was entered using pencils with high metal content. The cards were then read by an IBM reproducing punch equipped with a magnetic sensing unit. Since reproducing punches are still available at many schools, and since mark sense cards are easily produced, we developed the card shown in Figure 2. Students use both sides of the card. The card is then passed through an IBM reproducing punch twice, top down and top up, and punch reads the pencil marks and punches appropriate holes in the card. The punched mark sense card can then be read by computer and scored.

The system has the obvious disadvantage that it separates test taking from test scoring. The punch cannot be used in the classroom and several steps lie between testing and display of scores. Also, the IBM reader is not one hundred percent reliable. Reading in of cards is relatively slow and requires a trained operator. Yet, we have used the system for almost a year and found it acceptable. Schools that do not have optical scanners but have an IBM reproducing punch may want to consider this alternative.

Forms for Hewlett-Packard Optical Card Reader, Model HP7260A

We had the best experiences with this desk top card reader. The reader is relatively inexpensive, well engineered and easy to maintain. Only few malfunctions requiring outside service occured. Figure 3 shows FIT cards we developed for this reader. Students enter responses with a number 2 pencil. Cards are read singly or in batches, card images are transmitted by telephone to a computer, and the score is displayed immediately on a CRT. If a test consists of more than 12 questions, both sides of the card must be read by turning the cards over. Cards are filed away as protection against loss of records caused by computer malfunction. When printed in quantity, the cards cost about $3.50 per 1,000.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

31	DATE DAY	DATE MO	TEST NUMBER							SOCIAL SECURITY NUMBER									X
61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80

NAME

SIGNATURE

SS NUMBER

TEST NUMBER

DATE

COURSE

DS 2970 A

© OPTICAL SCANNING FORMS, INC. OPTICAL SCANNING CORPORATION

PRINTED IN U.S.A.

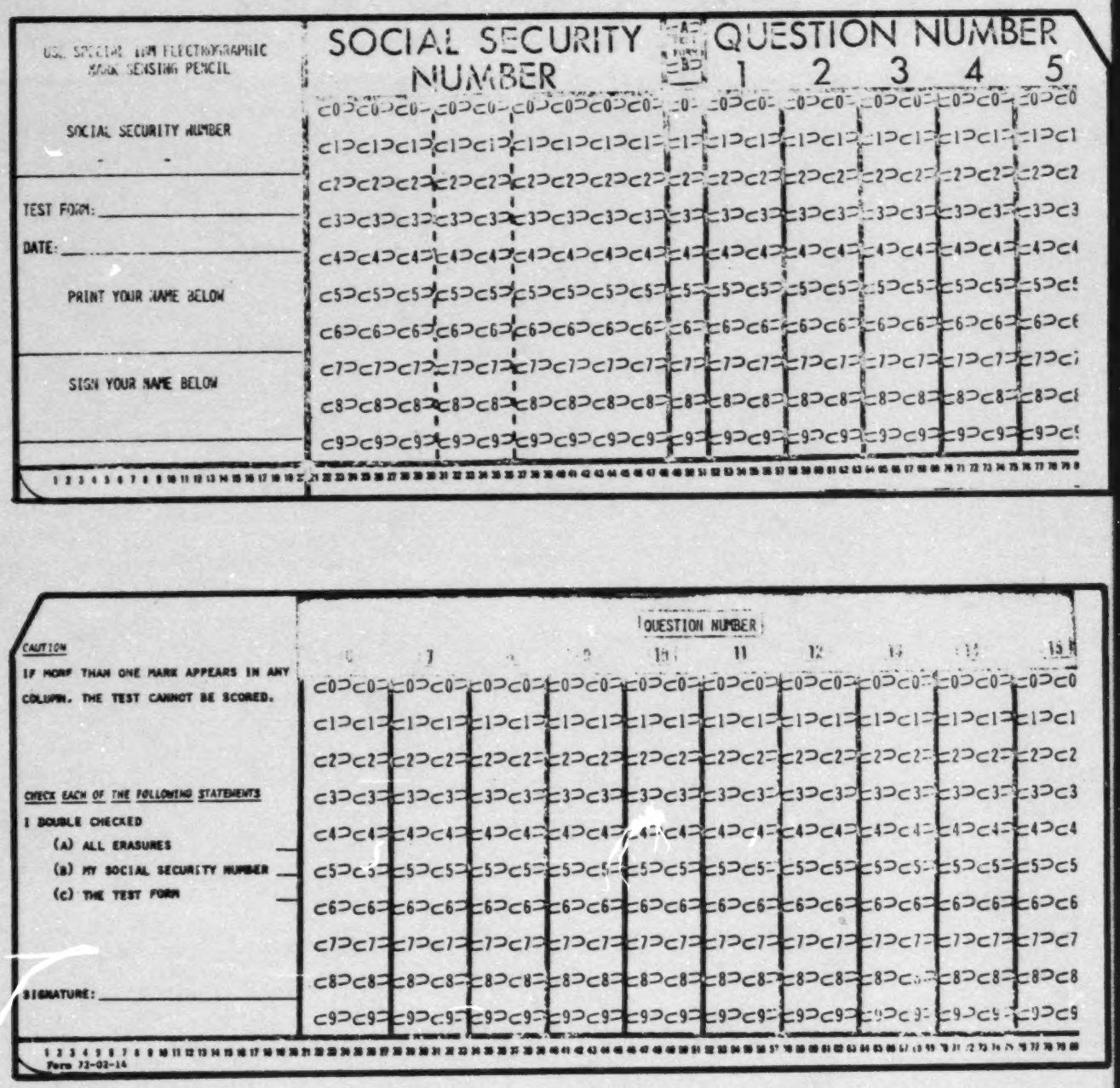
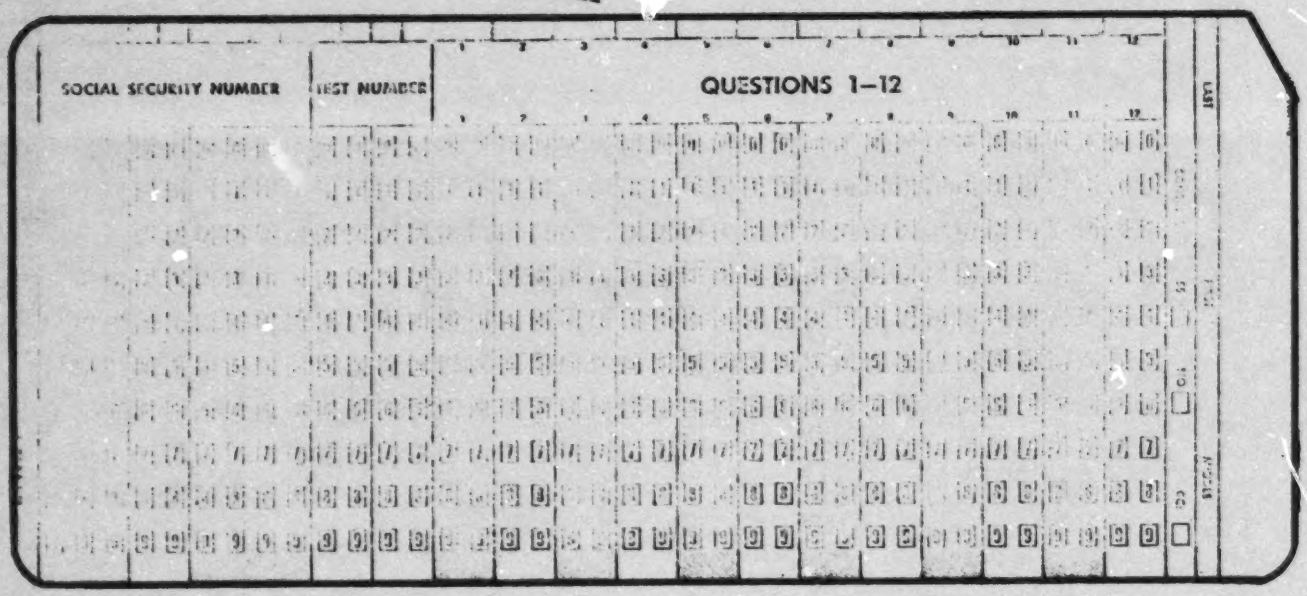


Figure 2. Front and back view of IBM mark sense card read and punched by IBM reproducing punch.

BEST COPY AVAILABLE



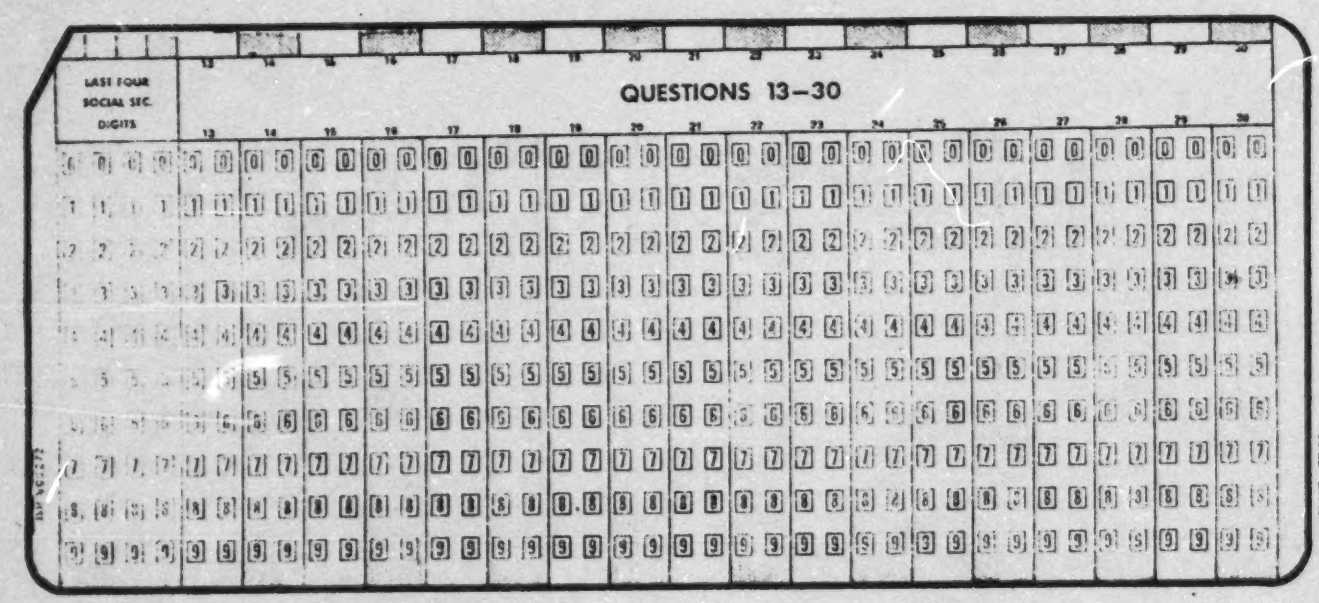


Figure 3. Front and back view of optical mark cards used with Hewlett-Packard card reader HP 7260 A.